CSL 356 Aug 1

Given a set S of n elements $x_1, x_2, \ldots x_n$, and an integer $1 \le K \le n$, we want to select an element in S with rank K.

$$\text{rank}(x, S) = \left| \{ x_i \in S \mid x_i < x \} \right|$$

$S = 6 \quad 3 \quad 9 \quad 4 \quad 20 \quad | \quad 3, 4, 6\ 9, 20$

$x = 3.8 \qquad \text{rank}(x, S) = 1$

Select (S, k) : returns an element in S with rank $= k$

min element : rank 1

Assume all elements S to be distinct

$[x_1, 1], (x_2, 2)\ (x_3, 3) \ldots \qquad (x_n, n)$

$x_i : x_j \qquad$ if $x_i < x_j \;\; \wedge \;\; x_j < x_i$

if $x_i = x_j \qquad$ smaller(i, j)

Sorting (S) vs. Selection(S, k)

(I) Selection is reducible to Sorting

$\propto$

(II) Sorting can be accomplished by multiple invocation of selection ✗

Selection(S, k) runs in $O(n \log n)$ comparisons

Can we select in $O(n)$ steps?

Suppose $\underline{k=1}$? or $k = n$ trivial

$k=2$, $k=3$

This procedure takes $O(k \cdot n)$ steps

$k = \frac{n}{2}$ (median) $\Omega(n^2)$

Look at the sorted set S (we don't sort)

$\tilde{x}_1, \tilde{x}_2, \tilde{x}_3 \quad \tilde{x}_{\frac{n}{2}} \; - \; - \quad \tilde{x}_n$

$\tilde{x}_i < \tilde{x}_{i+1}$

1. Choose an arbitrary element x from S

2. Lucky? Find rank(x, S)

Time n comparisons

What is the probability of success?

$= \frac{1}{n}$ using a random choice : every element is picked with equal probability

Pick up the k^{th} rank element.

Random variables X,

Expectation of X , $E[X]$

X = # times we iterate

$X \in \{1, 2, 3 \dots \}$

Probability distribution of X, say

$Prob[X = i] = p_i$

$$E[X] = \sum_{i \geq 1} i \cdot p_i$$

p_i follows geometric distribution

Fail $i-1$ times and succeed on the i^{th} trial where every trial is "independent"

$$p_i = ? \quad \left(1-\frac{1}{n}\right)^{i-1} \times \frac{1}{n}$$

If success prob. is p $\quad (1-p)^{i-1} \cdot p$

$$E[X] = ? \quad \frac{1}{p} = \frac{1}{\frac{1}{n}} = n$$

$$\Pr\left[X > k \cdot E[X]\right] \leq \frac{1}{k}$$

Markov's inequality for non-negative random variables

proof (by contradiction): Suppose j is the smallest integer such that $j > k \cdot E[X]$

Then $\sum_{t \geq j} t \cdot p_t > \sum_{t \geq j} j \cdot p_t = j \sum_{t \geq j} p_t$

$> k \cdot E[X] \left(\frac{1}{k}\right) > E[X]$ contradiction